CLIFFORD M. WITCHER

Some Communication Aspects
of Visual Prothesis. (1953)


EX
LIBRIS

Some Communication Aspects of Visual Prosthesis*

C. Lifford M. Witcher

Research Laboratory of Electronics
Massachusetts Institute of Technology
Cambridge, Massachusetts

Abstract

The communication system represented by the human eye, optic nerve, and visual cortex is capable of handling information at a rate far exceeding that of all our other sensory systems combined. For this reason it would seem, at present, completely impossible to devise for the visual system a mechanical substitute that could equal its performance under all conditions. However, in three important task areas, viz., reading of printed material, foot travel, and recognition of familiar objects at a distance, the information rates required are sufficiently small to permit the conception of suitable substitutes for vision. The basic problem in the design of a visual prosthesis for any or all of these areas is that of selecting from the environment the precise kinds of information necessary for the required task and the discovery of methods for encoding this information which will optimize the efficiency of the sensory channel used to convey it to the brain.

This paper, after reviewing a few notions connected with the theory of information, attempts to provide an analysis of the basic problem of visual prosthesis for two of the task areas mentioned above - reading and foot travel. Since the reading problem now appears to be solvable in terms of a device which will, within itself, be able to recognize and identify printed letters, the major emphasis is placed on the problem of mobility. The elements of a tentative solution to this problem are discussed, and descriptions of some devices embodying these elements are given.

*This work was supported in part by the Signal Corps, the Air Materiel Command, and the Office of Naval Research.

Some Communication Aspects of Visual Prosthesis

Blindness as an Informational Deficiency. Every sense organ, together with its sensory nerve and associated cortical area in the brain, constitutes a communication system. Conservatively, 95 per cent of our information about the external environment reaches us through the communication system composed of eye, optic nerve, and visual cortex. In fact, according to recent calculations by Jacobson,[1] the figure should be nearer 98 per cent. This is probably the best analytic conception of blindness to adopt as a point of departure when thinking about the problem of visual prosthesis. So far as it can be expected to go, a visual prosthetic device has to do the work of the eye while utilizing a quantity of information which is one-twentieth to one-fiftieth of that available to the eye.

It fortunately happens that in many visual tasks the brain does not need to utilize anywhere near the quantity of information which the eye is capable of pouring into it, and it is in connection with some of these tasks that visual prosthetic devices are considered feasible. Specifically, the three task areas in which work has been done are: enabling the blind to read printed or written material, enabling them to travel safely and independently, and attempting to provide recognition of familiar objects at a distance. The chief work has been done in the first two of these areas and will be considered in some detail below. Unfortunately, results achieved in the third are still so meager that they do not warrant discussion here.

Evidently, every sensory prosthetic device must constitute the input end of a communication system and as such it is the only part of the system over whose design we have any control. Once we have selected the sense modality which our device is to affect, the rest of the system (sensory nerve and cortical area) is predetermined within the limits of normal learning and habituation. It would then seem that in designing a prosthetic device, and most especially a visual prosthetic one, the basic problems involved are those of communication engineering and are of two specific kinds: (1) to select from all the information available from the environment precisely that which is relevant to the accomplishment of the desired task; and (2) to encode this information in order to utilize to its maximum efficiency the remainder of the communication system.

Remarks on Information. In what follows it will be assumed that the reader has at least a slight familiarity with the quantitative notion of information as a measure of the degree of arbitrariness, or freedom of choice, associated with a symbol's being transmitted through a communication system. Recall that the unit of information, or bit, represents the amount of information involved in a single binary choice. Letters chosen arbitrarily from a 26-letter alphabet are each said to carry an amount of information equal to $\log_2 26 = 4.7$ bits. A Braille symbol, in which 6 binary choices for presence or absence of dots are possible will, when chosen arbitrarily, carry 6 bits of information. This assumes that the blank space represents a possible symbol.

Recall that redundancy is a measure of the restriction of the freedom of choice of message symbols arising other than from the total number of different symbols available from a given source. Thus, in a language, the source of symbols is simply the letters of the alphabet, but the various inequalities in probability of occurrence of letters and similar unequal probabilities of word occurrence constitute redundancy. C. E. Shannon[2] has estimated that printed English is about 70 per cent redundant, so that, when occurring in printed text, a letter will represent on the average only about $0.3 \times 4.7 = 1.4$ bits of information. In so-called grade II Braille each symbol accounts for about 1.4 letters on the average; and thus, in an English text, each Braille symbol will carry about 2.0 bits of information. This means a redundancy for Braille of 4 bits per symbol, or 67 per cent.

Redundancy in any code utilized in a communication system is not always an objectionable characteristic. If the capacity of the communication channel is great enough, increased redundancy can make possible the detection of errors of coding and may even serve to correct errors. However, if the channel capacity is small and we wish to get the maximum rate of transmission of relevant information through it, redundancy has to be avoided. The writer wishes to emphasize his feeling that the point just made constitutes the basis on which communication theory can contribute most to the visual prosthesis problem.

Let us think, for a moment, of Braille symbols and of printed letters each as representative members of classes of patterns formed in accordance with the following general rules. In the case of Braille, let the rule be that the patterns occupy an area not greater than that of a single Braille cell and are composed of standard-sized Braille dots. (Experience has shown that these

are the conditions for best tactile discrimination.) For the ink patterns let us impose the same condition of approximately equal area and specify a width of ink line appropriate for best visual discrimination. It is precisely as patterns of these kinds that the impressions of Braille or ink symbols reach the brain of the reader. Evidently, in the case of the Braille patterns, we cannot actually conceive of many more possibilities other than the 64 standard Braille symbols (again including the blank space). However, in the case of the ink drawings, it is obvious that many hundreds, or even thousands, are conceivable. Thus, when thought of in this way, the redundancy of printed letters is enormously greater than that of Braille symbols. The visual system, with its high capacity for handling information, can often utilize this great redundancy of printed letters to make correct identifications, in spite of gross distortions in printing. On the other hand, the relatively low redundancy of Braille is one of the major factors in accounting for the fact that Braille reading speeds are one-fourth to one-half of that for reading by sight, even though the information is reaching the brain through, at most, a few score nerve fibers from the tip of the index finger.

Let us now see how these ideas may be applied to visual prosthesis.

The Prosthesis Problem for Reading. Any device for enabling the blind to read printed material must be a device which converts the visual patterns of the printed letters into patterns involving some other sense modality. During the 40 years or so in which work on this problem has been going on, two general types of devices have been conceived. The first kind, which we may call "direct translation" devices, converts the visual letter patterns into patterns of sound or touch which bear a rough resemblance in some one or more of their characteristics to the original letters. The second type, generally known as "recognition" devices, performs within itself the function of identifying the printed letters and thereafter produces predetermined forms of output patterns, such as pronunciation of the letters or presentation of their equivalents in Braille or Morse code.

The historic optophone and visagraph,[3] together with several later machines developed by the Haskins Laboratories and the Radio Corporation of America,[4] are examples of direct translation machines. None of these has proved practical, for reasons which will be considered presently. On the other hand, the so-called analyzing reader recently developed by David H. Shepard,[5] of the Intelligent

Machines Research Corporation, Arlington, Va., is a recognition type of machine that gives much promise of serving as a successful reader for the blind.

Shepard estimates that if a suitable variant of this machine were produced in relatively small quantities for use by the blind, the unit cost might be of the same order as that of an average home television receiver. The only area in which communication theory would seem capable of contributing to this type of machine would be in the selection of the best of the several standardized forms of output patterns and the study of the modes of presentation that yield maximum reading speed with minimum fatigue.

It is quite possible that a direct translation machine could be produced at a cost considerably below that of a recognition machine, and it is in the design of this form of machine that communication theory can contribute most. As we have already remarked, the output patterns from all direct translation machines, hitherto constructed, bear a rough resemblance to the input patterns. (This is not entirely true, since one attempt to get around this difficulty was made a few years ago by F. S. Cooper[3] of the Haskins Laboratories.) Hence these output patterns have all contained large amounts of redundancy. This statement, of course, applies to the magnified raised-letter patterns of the visagraph just as much as to the sound-generating machines. However, the visagraph patterns possessed one important property which the others did not. Since all parts of the letter produced by the visagraph could be touched simultaneously by the finger tip, the information from it reached the brain in an integrated form, just as in the case of Braille. It is for this reason that visagraph printing could definitely be read by most blind people, but it is the redundancy of this same printing which held reading speeds down, usually, to below 30 words per minute.

We have seen how Braille symbols, with their discrete presence-or-absence character, low redundancy, and tactile integrative nature, can provide remarkably high reading speeds. If we can design a direct translation reading machine whose output patterns likewise possess these three characteristics, its chances for practical utility will be greatly increased. In the case of sound outputs, the patterns might ideally be code groups of tones which were produced only after the machine had scanned and, so to speak, integrated an entire letter. These code groups might consist of two short tonal combinations, each of which is characterized by the presence or absence of 3 possible tones or tonal ranges. This sort

of presentation would be quite analogous to Braille, but could admit of some variability in the pitches of the tones arising from the variability in the printed letters.

The Mobility Problem. The past 10 years have witnessed much experimental work on travel aids for the blind and considerable philosophical evolution with regard to their necessary attributes. The nature of the problem is such that it does not permit of a solution as clear-cut as that for the reading problem. Even now the specifications for such a device are slightly fluid, but they can be roughly summarized as follows. (1) An adequate travel aid should enable the blind to avoid fixed obstacles of any size distributed with a maximum density of occurrence of 1-2 per second when a walking speed of 5 feet per second is maintained. (2) It should provide a positive, absolutely reliable warning of any step-down or sudden drop at a distance of 6-8 feet ahead, as well as permit search for step-downs on either side of the path of the user. (3) Its use should not afford any undue inconvenience to the user, such as excessive conspicuousness, need for a high degree of mental concentration, and the like. An almost equivalent way of putting this is to say that the user of a travel aid should be able to maintain a state of relaxation comparable to that of a person with normal sight when driving a car under average conditions. (4) Performance should be equally good under all normal weather conditions.

From these requirements, it will be seen that we do not, at the present, conceive a practical travel aid as a complete visual prosthesis for travel. In its present conceptual stage it would afford no protection against vehicular traffic. The fact that such moving obstacles as pedestrians, carts, etc. are not included in the obstacle avoidance requirement is not a limitation, since these obstacles produce their own warning signals so long as the blind person's ears are available to receive them. However, this imposes upon the device the restriction that its output patterns must be either nonauditory or at least of such a nature as to not interfere with the normal functioning of the ears.

A considerable number of obstacle-detection devices has been produced, both in this country and in England, but only one of them, the U. S. Signal Corps device, has been partially successful. A step-down detector is now being completed under the writer's direction at the Massachusetts Institute of Technology.

With but two exceptions* all obstacle detectors thus far produced have been of the narrow-beam or simple probe type; i.e., they were able to detect obstacles only in one direction relative to their own orientation. Thus, to obtain information about obstacles adjacent to either side of the user's path, a manual scanning technique had to be employed. None of these devices possessed any means for integration of their information. A variety of output signals, both audible and tactile, has been used.

Let us now examine the informational aspects of the mobility problem, considering, first, step-down detection, since it is the simplest aspect from the communication standpoint. Some years ago it was realized that step-downs represented the greatest hazards to independent travel by the blind. It was soon found that if the U.S. Signal Corps device was held so that it "looks at" a spot on the ground a few feet ahead, a step-down of sufficient height could be detected by virtue of the change in range which occurred when the device "looked" over its edge. The difficulty was that when such a device was used as a step-down detector, the user had to remain continually attentive to the output signal in order to be sure to catch the momentary change in range which signified a step-down. Thus the signal represented redundant information continuously being passed through the sensory channel connecting the device with the brain of the user. When work began on the M.I.T. step-down detector, we postulated that a satisfactory device would have to be one in which the output is zero until the user approaches to within a prescribed distance of a step-down (6-8 feet), whereupon a definite, relatively intense alarm signal would be produced.

Let us turn now to the obstacle detection problem. In 1949 the writer,[8] working with a group of experienced blind people in New York known as the Technical Research Council, proposed the following analysis of the information requirements for an obstacle detector. Let the space in front of the user be divided into three regions as in Fig. 1. The outer region, R_1, we may call the "awareness region." Let us call the middle area, R_2, the "attention region,"

*The beam from the British clicker device,[6] which generated high-frequency sound pulses, was deliberately made somewhat divergent with the idea that the user would be able to hear the reflections of these pulses from obstacles near each side of his path. The pattern optical device[7] constructed at Haskins Laboratories, New York City, represented an attempt to provide recognition of simple objects at a distance. It utilized automatic scanning.

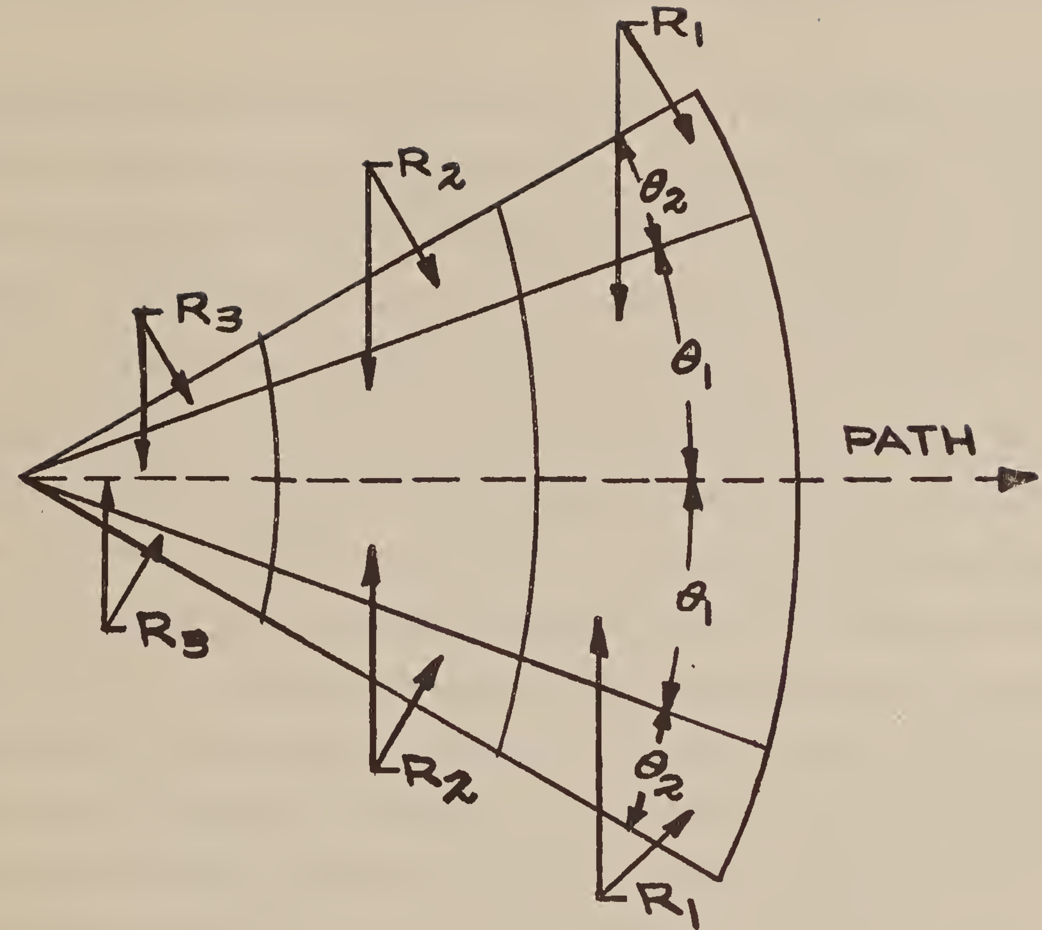


FIG. 1

and designate R_3 as the "avoidance region." If an obstacle enters R_1, it would be desirable, though not absolutely necessary, for the user to be made aware of the fact. Once an obstacle passes from R_1 into R_2, its presence must be immediately known by the user, and he must devote sufficient attention to it to arrive at a decision as to whether or not he must change his course in order to avoid it. If avoidance action has to be taken, it must be initiated by the time the obstacle passes from R_2 to R_3.

On the basis of reaction-time considerations and with the assumption of a walking speed of 5 feet per second, a length of 4-5 feet was estimated for both R_2 and R_3. The length of R_1 may conveniently be taken equal to that of the other regions. This was the extent of the original analysis, but now let us carry it a bit further.

The important region from which information must be picked up by the detector and transmitted to the user's brain is R_2. An ideal detector should provide the brain continuously with the answers to three questions of the binary-choice type: (1) Is there, or is there not, an obstacle present in R_2? (2) If an obstacle is present, is it predominantly on the right or left side of the path?* (3) Is the obstacle close enough to the path to necessitate a change of course to avoid colliding with it? Thus we see that for each obstacle considered 3 bits of information must be made available as soon as possible by the detector. Assuming a length of 5 feet for R_2, an information rate of 3 bits per second per obstacle is thus needed, with a possible extra bit for indication of presence or absence of an obstacle in the awareness region. On the basis of what we have already noted for the case of Braille reading, it should be evident that this rate of flow of information can be easily handled by the tactile channel from even one finger tip, provided the information is properly encoded and is to some extent integrated.

Referring again to Fig. 1, the angles θ_1 represent the limits within which the presence of an obstacle in R_2 definitely requires avoidance action. Assuming a body width of 20 inches and a possible deviation from a straight course of 1 foot in 5 feet of travel, $\theta_1 = \tan^{-1} 22/60 = 20°10'$. One might ask why the detector should pick up information for angles greater than θ_1. The answer is that

*If the obstacle extends completely across the path, exploratory rather than avoidance action must be taken.

the knowledge that an avoidance action has shifted the position of an obstacle into one of the outer angular regions θ_2 provides positive confirmation that the avoidance movement has been sufficient.

The analysis given here necessarily implies a detector utilizing automatic scanning with an angular amplitude equal to $\theta_1 + \theta_2$ and a frequency which is large compared to V/R_2, where V is the walking speed. (θ_2 can be conveniently taken as about 10°.)

The first scheme for presenting spatially integrated obstacle-detector information was proposed by the writer about 3 years ago and experimentally investigated by N. O. Sokal[9] at the Massachusetts Institute of Technology. In this form of presentation the whole region of Fig. 1 was mapped onto a small similarly shaped area on a signal presentation plate. This area contained an array of holes through which small pins could project to indicate the presence of an obstacle. Since the presentation region was small enough to be scanned rapidly by the user's fingers, he had available at all times a rough PPI representation of the distribution of obstacles in the region covered by the detector. This output scheme was not actually applied to an obstacle detector, but a simulator arrangement was used. Tests with several blind subjects indicated that the scheme was capable of feeding information to the brain considerably faster than was possible with any existing obstacle detector. The major difficulty with the scheme lies in the engineering problems connected with its incorporation into an obstacle detector.

A simpler scheme, suggested by Prof. T. A. Benham,[10] would seem to be more practical at this stage of the game. Benham proposes to present the information tactually through a small number (4 or 5) of stimulus points on the handle of the device. Each of these points would correspond to a different angular region, so that the scheme partially utilizes spatial integration. Range would be indicated by some signal variable associated with the stimulus points, such as vibration rate. Alternatively, the stimulator points might simply be made to press against the fingers continuously so long as an obstacle is present in the attention region. With this variant of the scheme, the 2 middle points of a set of 4 would indicate right or left orientation of the obstacle within the angular limits θ_1, and the 2 outer points would indicate the presence of an obstacle in either of the regions θ_2. Using a weak, perhaps relatively high frequency, vibration of the entire handle to indicate an obstacle in region R_1, and a relatively strong low frequency handle

vibration to indicate a step-down, the handle of the detector would apparently be providing adequate information for safe travel in the sense already defined.

A handle and signal simulator incorporating this scheme have recently been built at the Massachusetts Institute of Technology, and field tests with these devices are now in progress.

References

(1) H. Jacobson, "The Informational Capacity of the Human Eye," Science 113 (1951) pp. 292-293

(2) C. E. Shannon, "Prediction and Entropy of Printed English," B.S.T.J. 30 (Jan. 1951) pp. 50-64

(3) F. S. Cooper, "Research on Reading Machines for the Blind," from Blindness, P. A. Zahl, ed., Princeton University Press (1950) pp. 512-543

(4) V. K. Zworykin, L. E. Flory, and W. S. Pike, "Research on Reading Aids for the Blind," J. Fr. Inst. 247 (May 1949) pp. 483-496

(5) D. H. Shepard, "The Analyzing Reader," paper delivered at Boston meeting of Association for Computing Machinery (Sept. 1953) not yet published

(6) R. L. Beurle, "Electronic Guiding Aids for Blind People," Electronic Eng. XXIII (Jan. 1951) pp. 2-7

(7) Appendix T of "Research on Guidance Devices and Reading Machines for the Blind," Final Report of Haskins Laboratories (1947) pp. T1-T23

(8) C. M. Witcher, "General Considerations on Guidance Devices," paper prepared for Prosthetic and Sensory Aid Service, Veterans Administration, New York (March 1949)

(9) N. O. Sokal, "Area Display of Obstacle Location for Use with a Guidance Device for the Blind," Master's thesis, Department of Electrical Engineering, M.I.T. (Sept. 1950)

(10) T. A. Benham, oral communication with the writer (Dec. 1952)

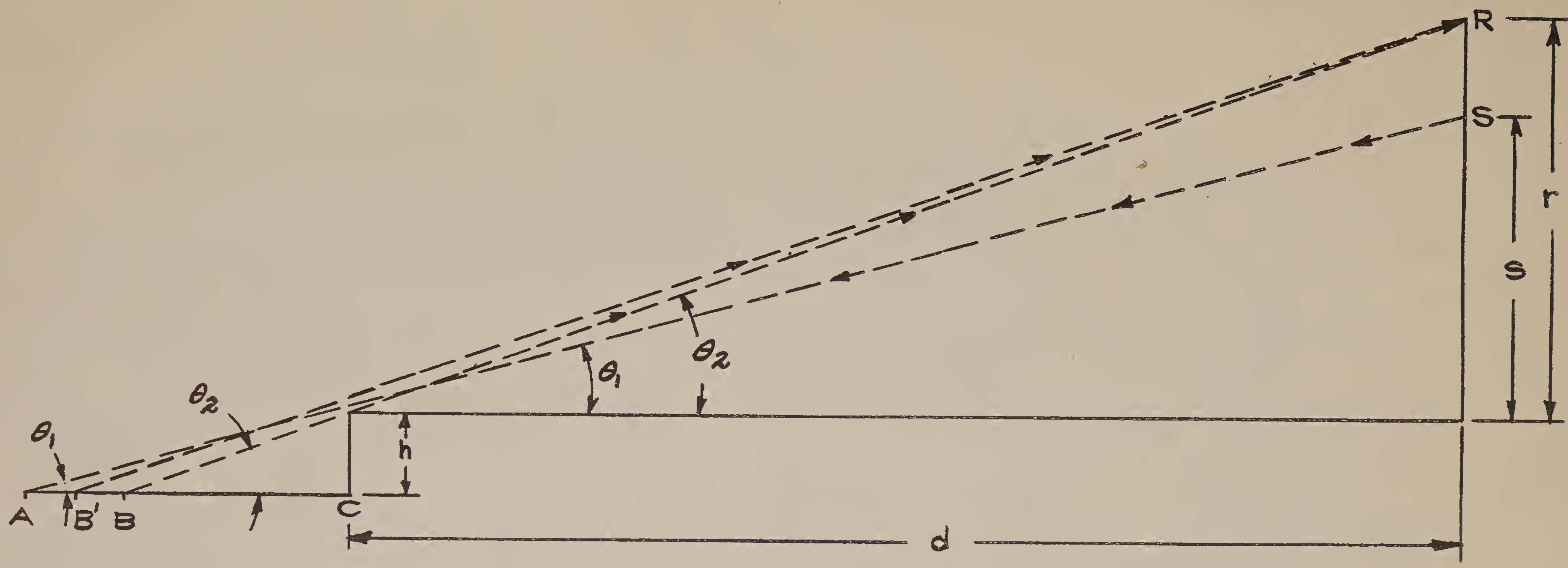
Appendix 1: The M.I.T. Step-Down Detector

The basic components of this device are a vibratory light source, an optical receiving system, an amplifier, and an alarm circuit. The relative positions of the source, S, and receiver, R, can be seen by reference to Fig. 2. The source (comprised of a flashlight bulb and reflector) projects a beam of light forward and downward toward the ground. By virtue of the fact that the source executes angular oscillations about an axis perpendicular to the plane of the paper, the light beam swings back and forth in the plane of the paper, striking the ground along a line extending from about 5-9 feet in front of the device when the latter is held at the height shown.

The receiving system, consisting of a mirror, lens, and photo cell, is sharply focused so that it only picks up light from a narrow spot on the ground, indicated by BB' in the figure. If the step-down, h, of the figure is not present, it should be apparent that the receiver must pick up a pulse of light each time the beam from the source swings across the image space BB'. Thus, when the device is "looking at" level ground, the amplifier and output circuit are receiving a continuous succession of signal pulses. However, if a step-down is present at the position shown in the figure, it is apparent that (except for slight diffraction effects) no light from the source can fall inside the region AC, and hence no signal will be picked up by the receiver. The alarm circuit at the output end of the amplifier is so arranged that, if a single pulse in the regular succession of signal pulses fails to appear, it will set off a vibratory alarm in the handle of the device. Thus the alarm provides a positive indication of a step-down at a distance of 6-8 feet in front of the device.

The constants of the device are such as to insure the operation of the alarm at speeds of approach to the step-down as high as 15 feet per second, which is much faster than normal walking speed. The scanning motion of the light beam produced by the oscillation of the source serves two useful purposes: (1) it provides a pulsed signal, which is the most desirable kind for the present case; and (2) it insures that a moderate amount of swing of the device as it is held in the hand while walking will not cause the signal to disappear momentarily and thus produce false alarms.

Ideally, the presence of a step-down is the only condition which can cause the alarm to operate. However, when the surface of the ground is covered with ice or water, much of the light from the source is lost through specular



$r = 28''$
$s = 21''$
$d = 84''$
$\tan \theta_1 = 0.25$
$\tan \theta_2 = 0.33$
$BC = 3h$
$AC = 4h$

FIG. 2

OPTICS OF STEP-DOWN DETECTION

reflection, so that the signal tends to disappear. This difficulty has now been overcome, to some extent, by greatly increasing the gain of the receiving amplifier, but the present device will still give false alarms when the image space BB', falls on the surface of a puddle with a black dirt bottom. There is still also some tendency of the device to give false alarms produced by blocking of the amplifier when the image space moves from bright sunlight into shadow or vice versa.

Note:

For a more detailed discussion of the evolution and operating characteristics of this device, the reader is referred to "Replacement of Visual Sense in Task of Obstacle Avoidance" by C. M. Witcher and L. Washington, Jr., Quarterly Progress Report, Research Laboratory of Electronics, M.I.T., July 15, 1952, et seq.

Date Due			
~~MAY 30 '68~~			